DOCUMENT RESUME

ED 346 747 FL 020 490

AUTHOR	Pierce, Lorraine Valdez; O'Malley, J. Michael
TITLE	Performance and Portfolio Assessment for Language Minority Students. Program Information Guide Series, 9.
INSTITUTION	National Clearinghouse for Bilingual Education, Washington, DC.
SPONS AGENCY	Office of Bilingual Education and Minority Languages Affairs (ED), Washington, DC.
REPORT NO	NCBE-9
PUB DATE	92
CONTRACT	T289004001
NOTE	34p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Bilingual Students; Check Lists; Data Collection; Evaluation Criteria; *Evaluation Methods; Holistic Approach; Informal Assessment; Intermediate Grades; Junior High Schools; Language Proficiency; *Language Tests; *Limited English Speaking; Minority Groups; *Portfolios (Background Materials); Profiles; *Student Evaluation; *Student Records

ABSTRACT

Performance assessment procedures and a portfolio assessment framework are presented for monitoring the language development of language minority students in the upper elementary and middle grades. Unlike standardized achievement tests, performance-based assessment can effectively monitor the progress of language minority students because it can be conducted on an ongoing basis. Portfolio assessment has the potential for becoming an effective management system for performance assessment results. It can be conducted on a classroom, school-wide, or district-wide basis. This report contains two main sections. The first discusses performance assessment, covering purposes, types, design, administration, and scoring techniques for oral language assessment, reading assessment, writing assessment, anecdotal records, and self-assessment for language proficiency and learning strategies. A sample literacy development checklist and a chart showing sample holistic criteria are provided. The second section reports on portfolio assessment, including the following: designing portfolios (e.g, purpose, specific focus, performance task/instrument selection, setting criteria, staff collaboration, staff development, student selection, and student/parent involvement); planning for and collecting the data; analyzing portfolio contents; using portfolio results; and some common concerns about putting portfolio use into practice (e.g., time requirements, content interpretation, ways to use portfolios and performance-based assessment). Sample contents from reading/writing portfolios in Virginia, New Hampshire, and Florida are summarized in a one-page chart. Contains 68 references. (LB)

ED346747

NCBE

9

Performance and Portfolio Assessment for Language Minority Students

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

by

[redacted]

Lorraine Valdez Pierce

J. Michael O'Malley

National Clearinghouse for Bilingual Education

Program Information Guide Series

Spring 1992

BEST COPY AVAILABLE

FL 020 490

Acknowledgments

The authors would like to thank the following staff members of the NCBE for their comments on the manuscript for this Program Information Guide: Denise McKeon, Holly Fadden, Minerva Gorena, Joel Gomez, and Donna Christian. In addition, we would like to thank our external readers for their thoughtful reviews: Yvonne and David Freeman, Fresno Pacific College; Margo Gottlieb, Illinois Resource Center; and Deborah Short, Center for Applied Linguistics.

The National Clearinghouse for Bilingual Education (NCBE) is funded by the U.S. Department of Education's Office of Bilingual Education and Minority Languages Affairs (OBEMLA) and is operated under Contract No. T289004001 by The George Washington University's Center for the Study of Education and National Development, jointly with the Center for Applied Linguistics. The contents of this publication do not necessarily reflect the views or policies of the Department of Education, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. Readers are free to duplicate and use these materials in keeping with accepted publication standards. NCBE requests that proper credit be given in the event of reproduction.

National Clearinghouse for Bilingual Education
1118 22nd Street NW, Washington DC 20037

Director: Joel Gómez
Publications Coordinator: Lorraine Valdez Pierce

Introduction

Well-designed assessment procedures are essential to meeting the needs of language minority students acquiring English as their second language. Assessment is involved at many steps in a continuum of services for these students: in initial identification, in the placement of students into appropriate instructional programs, in monitoring the progress students make within these programs, in reassigning students to different levels within a program depending on their growth in English language skills, in moving students out of special programs and into mainstream classes, and in following the progress of these students in the mainstream. This continuum is wholly dependent at each step on the appropriate selection, use, and interpretation of relatively complex assessment procedures.

Background

Recently, there has been a growing interest among mainstream educators in performance assessment due to concerns that multiple-choice tests, usually the only option available from test publishers, fail to assess higher order skills and other skills essential for functioning in school or work settings (Haney & Madaus, 1989; Neill & Medina, 1989; O'Neil, 1992; Wiggins, 1989). Multiple-choice tests are not authentic because they do not represent activities students typically perform in classrooms. In addition, multiple-choice tests do not reflect current theories of learning and cognition and are not based on abilities students actually need for future success (Herman, 1992). Another concern is that standardized tests cannot be used to closely monitor student progress in the school curriculum throughout the year since they are only administered once or twice annually. These concerns are no less valid for educators of language minority students.

Purpose and Definitions

This publication describes performance assessment procedures and a portfolio assessment framework for monitoring the language development of language minority students in the upper elementary and middle grades. Performance and portfolio assessment may also be used to meet other purposes, such as reassignment or reclassification, as determined by teachers or other school staff. Although assessment of student progress in the content areas merits description and critical analysis, examples of performance assessment procedures provided here are limited to the monitoring of English and/or native language development. We begin by defining basic terms to be used throughout this publication: alternative assessment, performance assessment, and portfolio assessment (Baron, 1992a; Stiggins, 1987; Tierney, Carter, & Desai, 1991).

Alternative assessment:

- is any method of finding out what a student knows or can do that is intended to show growth and inform instruction and is not a standardized or traditional test;
- is by definition criterion-referenced;
- is authentic because it is based on activities that represent actual progress toward instructional goals and reflect tasks typical of classrooms and real-life settings;
- requires integration of language skills; and
- may include teacher observation, performance assessment, and student self-assessment.

Performance assessment:

- is a type of alternative assessment;
- is an exercise in which a student demonstrates specific skills and competencies in relation to a continuum of agreed upon standards of proficiency or excellence; and
- reflects student performance on instructional tasks and relies on professional rater judgment in its design and interpretation.

Portfolio assessment:

- is the use of records of a student's work over time and in a variety of modes to show the depth, breadth, and development of the student's abilities;
- is the purposeful and systematic collection of student work that reflects accomplishment relative to specific instructional goals or objectives;
- can be used as an approach for combining the information from both alternative and standardized assessments; and
- has as key elements student reflection and self-monitoring.

Performance assessment and portfolios are complementary approaches for reviewing student language development and academic progress. Together they represent authentic assessment, continuous assessment of student progress, possibilities for integrating assessment with instruction, assessment of learning processes and higher-order thinking skills, and a collaborative approach to assessment that enables teachers and students to interact in the teaching/learning process. One of the advantages of using performance and portfolio assessment with language minority students is that the assessment can be conducted in the students' native language (O'Malley & Pierce, 1991).

Portfolios have been most widely used in the teaching of reading and writing, with a strong focus on classroom instruction, student ownership and self-evaluation, and teacher autonomy (Graves, 1983; Tierney, Carter, & Desai, 1991). More recently, portfolios have been proposed and adopted as statewide student assessment management tools in addition to or instead of standardized achievement test data (Baron, 1992b; Brewer, 1990; O'Neil, 1992; Rothman, 1991; Vermont State Department of Education, 1990). Provided that the purpose of portfolio contents is to expand understanding of a student's growth based on multiple measures, different kinds of test and non-test data can be included in a portfolio. Portfolios might also contain required information for state- or district-wide systems, but these data need not dominate or divert portfolio assessment from being used to inform classroom instruction.

Performance assessment and portfolios are typically seen as sources of teacher and student empowerment because control over assessment shifts from the administrators to those linked most closely to instruction (Graves, 1983; Tierney, Carter, & Desai, 1991). While we believe this shift of control over assessment is a positive one, we also believe that the most useful kind of assessment is that which can be shared with other teachers who interact or will interact with students. For this reason, we emphasize systematic approaches to assessment which can ensure the reliability and validity of the results.

Performance Assessment

Performance assessment falls into two categories: achievement-related behaviors exhibited by the student (e.g., communication or reading skills) and achievement-related products that students develop (e.g., written reports or projects). Performance assessments require that the assessor:

(1) observe the behavior or examine the product that is reflective of achievement, and

(2) apply clearly articulated performance criteria so as to make a sound professional judgment regarding the level of proficiency demonstrated. Intuitions, impressions, and "feelings" about student performance are *not* a part of sound performance assessments. Like paper-and-pencil tests, performance assessments must adhere to certain rules of evidence (Stiggins, 1990).

In this section, procedures are provided for the design of performance assessment tasks and instruments that can be included in portfolios of language minority students and children learning English as a non-native language. Procedures for assessing specific language skill areas (oral language, reading, and writing) and their integration are described. Suggestions are also provided for designing student self-assessment measures of language proficiency and language learning strategies. Each category below includes brief descriptions of types of assessment procedures and ways to design and administer performance tasks, with particular attention to the development of appropriate scoring procedures. Each assessment technique should be adapted as needed to match the developmental and language proficiency levels of the students. This can include assessment in the students' native language.

Oral Language Assessment

Purpose/Types

To determine oral language comprehension and production, teachers can administer performance assessments which reflect tasks typical of the classroom or real-life settings. In this way, assessment is authentic and aligned with both the curriculum and students' prior experience. Oral performance assessments are not limited to a single type and can take various forms depending on their authenticity in relation to classroom activities. These can include: oral interviews, story retellings, simulations/situations, directed dialogues, incomplete story/topic prompts which students are asked to complete, picture cues, teacher observation checklists, and student self-evaluations (Bachman & Palmer, 1989; Gonzalez Pino, 1988; Omaggio, 1986; Oscarson, 1989).

Design/Administration

A major obstacle to conducting oral language assessment in the classroom is the time involved in administering the assessment to students one at a time. Alternatives to individual student assessment in large classes include the use of teacher observation checklists and the assessment of pairs or small groups of students at one time (Fox & Allen, 1983; Genishi, 1985; Gonzalez Pino, 1988; Pinnell, 1985).

To conduct assessment in pairs, teachers can use *cue cards* and ask students to interview each other or to interact following the directions on the cue cards (Gonzalez Pino, 1988). Cue cards should be written at the reading level of the students to be assessed. In order to have enough oral language production to assign a rating, at least five or six sentences should be

elicited from each student. Cue cards can easily be constructed by providing written directions, called *prompts*, on index cards to elicit the performance teachers wish to assess. For example, if a lesson has just been conducted on personal greetings and leave-takings, one student's cue card might read, "Greet your classmate and find out where he or she has been for the past few days." The other cue card could read, "Respond to your classmate's questions by telling him or her that you have been at home sick. Find out from your classmate what you have missed in class."

Depending upon the age and language proficiency levels of students in the class, teachers can use *picture cues* or *topic monologues* to conduct individual oral assessments. To use picture cues, the teacher presents pictures or photographs appropriate for the age and interest level of the students being assessed. From among several pictures the teacher presents, students can be asked to choose one or two pictures that they feel they can talk about. The teacher can lead the student into talking by asking questions such as "What story does this picture tell? Has this ever happened to you? What do you think will happen next?"

In using topic monologues, the teacher provides a topic for the student to address at length. This is an authentic performance task for upper elementary and middle school students who are often asked to present oral reports and syntheses of what they have read. If a lesson has just been presented on foods to students at an intermediate level of English proficiency, for example, a topic monologue such as the following might be appropriate: "Tell me about your favorite food. Describe it and tell me why it is your favorite." For more advanced students, the topic might be: "Give me the recipe for your favorite food."

Story retelling is an effective way to integrate oral and written language skills. Students who have just read or listened to a story might be asked to retell the story or to summarize the main idea. When pictures, topic monologues, cue cards, or story retelling techniques are used, instead of the typical interview pattern of teacher-question/student-response/teacher-question, teacher talk is reduced and more time is allowed for student language production.

Teachers may also want to assess students' ability to use academic language. Academic language is the cognitively demanding and contextually-reduced language of content area instruction and is critical for success in mainstream classrooms (Cummins, 1982; 1983). Academic language functions are the essential communication tasks that students must be able to perform in different content areas; they determine whether the learning task will be simple or complex (Dix, 1992; O'Malley & Chamot, 1990; Pinnell, 1985). For that purpose, teachers can identify specific language functions that are germane to an instructional goal or activity (O'Malley, 1991), such as:

- *Seeking information*

  Using language to explore the environment or acquire information (e.g., using *who, what, when, where,* and *how* to collect information);

- *Informing*

  Reporting, explaining, or describing information or procedures (e.g., retelling a story, telling the main idea of a reading passage, summarizing);

- *Analyzing*

  Separating a whole into its parts (e.g., telling parts or features of a concept, a procedure, an object); and

•*Evaluating*

Assessing the worth of an object, opinion, or decision (e.g., selecting or naming criteria to evaluate, prioritizing a list and explaining criteria, indicating reasons for agreeing or disagreeing).

Scoring

Teachers and/or raters should establish scoring criteria for a range of grade levels (e.g., Grades 4–6, 7–9) and identify at least three categories of proficiency: non-English proficient (NEP), limited English proficient (LEP), and fluent English proficient (FEP). Because these categories were originally intended for identification and placement purposes and consequently tend to be broadly defined, for the purposes of monitoring student language development it may be useful to differentiate bands of proficiency within each level, such as low, intermediate, and high. This will enable the teacher to design instruction more appropriate to students' needs and to monitor growth from one band to the next within levels as well as from level to level.

Scoring criteria should be holistic, with a focus on the student's ability to receive and convey meaning. Holistic scoring procedures evaluate performance as a whole rather than by its separate linguistic or grammatical features. Depending on the goals of instruction, grammatical accuracy and pronunciation may not be as important as skills critical to comprehending and conveying meaning. Students are evaluated by using detailed criteria or definitions of performance matched to a rating scale (Hamayan, et al., 1985; Navarrete, et al., 1990).

Teachers wanting to score grammar and pronunciation may wish to treat these as subscores and assign them weights which are less than the weight assigned to a subscore for overall communicative ability (Hamayan, 1985; Gonzalez Pino, 1988). Well in advance of the oral assessment, students should be provided with an explanation of how they will be rated. If teachers plan to assign a grade to the oral assessment, they will have to determine the scoring range which corresponds to specific letter grades at each grade level.

A holistic oral proficiency rating scale that teachers may want to use is the Student Oral Proficiency Rating (SOPR). The SOPR is a matrix (adapted from the Student Oral Language Observation Matrix/SOLOM developed by the San Jose Unified School District, California) which allows for rating students in five categories of oral language proficiency: comprehension, fluency, vocabulary, pronunciation, and grammar (see **Figure 1**). Within each category, students can be rated at five different proficiency levels. The ratings for each category are considered separate sub-scores, each ranging from 1 to 5, with 5 indicating the approximate level of proficiency for a native speaker. A total score results when sub-scores for the five categories are combined (Development Associates, 1988).

Reading Assessment

Performance assessment of reading with students who are learning English should focus on reading comprehension rather than isolated reading skills, as is typical of many standardized achievement tests. We suggest four approaches for the assessment of reading that have been amply described in the literature: miscue analysis (Goodman & Burke, 1972); the individual reading inventory (Cunningham, et al., 1983); anecdotal records (Goodman, Goodman, & Hood, 1989; Rhodes & Nathenson-Mejia, 1992); and cloze tests (Cohen, 1980; Hamayan, et al., 1985; Madsen, 1983; Oller, 1979). One additional approach which has not been adequately described in relation to the reading behaviors of language minority students and which we describe below is the use of rating scales.

Figure 1. Student Oral Proficiency Rating*

Student's Name	Grade	Language Observed	TOTAL SCORE
School		City	State
Rated by	Date		

DIRECTIONS: For each of the 5 categories below at the left, mark an "X" across the box that best describes the student's abilities.

CATEGORY	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4	LEVEL 5
A. Comprehension	Cannot understand even simple conversation.	Has great difficulty following what is said. Can comprehend only "social conversation" spoken slowly and with frequent repetitions.	Understands most of what is said at slower-than-normal speed with repetitions.	Understands nearly everything at normal speed, although occasional repetition may be necessary.	Understands everyday conversation and normal classroom discussions without difficulty.
B. Fluency	Speech is so halting and fragmentary as to make conversation virtually impossible.	Usually hesitant; often forced into silence by language limitations.	Speech in everyday communication and classroom discussion is frequently disrupted by the student's search for the correct manner of expression.	Speech in everyday communication and classrom discussion is generally fluent, with occasional lapses while the student searches for the correct manner of expression.	Speech in eveyday conversation and in classroom discussions is flueni and effortless, appoximating that of a native speaker.
C. Vocabulary	Vocabulary limitations are so extreme as to make conversation virtually impossible.	Misuse of words and very limited vocabulary make comprehension quite difficult.	Frequently uses the wrong words; conversation somewhat limited because of inadequate vocabulary.	Occasionally uses inappropriate terms or must rephrase ideas because of inadequate vocabulary.	Use of vocabulary and idioms approximates that of a native speaker.
D. Pronunciation	Pronunciation problems so severe as to make speech virtually unintelligible.	Very hard to understand because of pronunciation problems. Must frequently repeat in order to be understood.	Pronunciation problems necessitate concentration on the part of the listener and occasionally lead to misunderstanding.	Always intelligible, though one is conscious of a definite accent and occasional inappropriate intonation patterns.	Pronunciation and intonation approximate a native speaker's.
E. Grammar	Errors in grammar and word order so severe as to make speech virtually unintelligible.	Grammar and word order errors make comprehension difficult. Must often rephrase or restrict what is said to basic patterns.	Makes frequent errors of grammar and word order which occasionally obscure meaning.	Occasionally makes grammatical or word order errors which do not obscure meaning.	Grammatical usage and word order approximate a native speaker's.

***This form is an adaptation of the Student Oral Language Observation Matrix (SOLOM) developed by the San Jose (California) Unified School District.**

Adapted by Development Associates, Inc., Arlington, Virginia, and used with permission. **Signature of rater** _____

Purpose

Rating scales can be used to collect information on a range of student reading behaviors. Rating scales offer several advantages: they are systematic, require little teacher time, and do not disrupt instructional time. A rating scale is a checklist that contains predetermined performance tasks and a systematic procedure for indicating a student's level of performance relative to specific criteria. The number of levels in the criterion will be determined by the focus of the assessment, but each will be precisely defined.

Design

Four types of information teachers can maintain in using a rating scale for reading comprehension are reading skills, interest, applications, and reading strategies. Teachers can collect information as frequently as is considered appropriate to the purposes of assessment and instruction. For example, occasional checks on student progress could be conducted biweekly, monthly, or quarterly, although other options are also possible.

The assessment of *reading skills* is designed to reflect the student's ability to perform functional reading tasks. Although teachers can select their own objectives depending on local curriculum frameworks in language arts, some possible objectives are reflected in **Figure 2**. Comprehension of stories read aloud can be expanded to include literal and inferential comprehension, if desired. Teachers can determine if students have fluent decoding skills by asking them to read aloud individually and by checking comprehension through probe questions, such as "What was the main idea?" and "What do you predict will occur next?"

One of the major concerns in reading assessment should be a student's *level of interest* in reading materials. This interest can be shown in the variety of materials independently selected by students. The *applications* category reflects integration of speaking, reading, and writing skills. Applications may include a project, such as writing and presenting a skit, producing artwork, writing a paper, keeping a reading log on the kinds of materials read and how often they are read, or producing a self-report on reading ability. A teacher narrative or rating scale could be used to assess student progress in each of the categories listed in Figure 2. These categories show what students can actually do with reading materials and how they use them to achieve their own purposes.

A fourth possible category for rating scales is the students' use of *reading strategies*. Substantial evidence indicates that students who use reading strategies comprehend text more effectively than students who do not (e.g., Dole, Duffy, Roehler, & Pearson, 1991; Pressley, 1990), and that students from language minority backgrounds also benefit from the use of these strategies (O'Malley & Chamot, 1990; Padron & Waxman, 1988). Used in conjunction with miscue analysis, reading strategies assessment can reveal what kinds of strategies, techniques, or approaches students use to understand and remember information they have read, what they do when they are given a reading assignment, and what they do to understand and remember what they have read upon completing a reading passage. Some sample questions for determining reading strategies (modified from Goodman, Watson, & Burke, 1987) are:

1. When you are reading and come to something you don't know, what do you do?
2. Who is a good reader you know? What makes that person a good reader?
3. If you knew someone was having trouble reading, how would you help that person? What would your teacher do to help that person?
4. What would you like to do better as a reader?

From these kinds of questions, and from probe questions asked during an individual reading inventory, teachers may be able to rate student performance relative to the strategies indicated in Figure 2.

Administration

Teachers can collect information on reading skills using a rating scale while observing students individually, in pairs, or in small groups. Ratings for specific skills such as literal and inferential comprehension could be based on probe questions asked of students following a silent reading exercise. Probe questions should reflect different levels of cognitive complexity, such as the following (Maryland State Department of Education, n.d.):

•*Knowledge*

Who was the main person in this text? Where did the event take place?

•*Comprehension*

What was the main idea of the reading? Retell what you have read in your own words.

•*Analysis*

What are the parts or features of _____? How does _____ compare with _____? Outline the components of _____. What evidence supports _____?

•*Synthesis*

What would you predict will occur next? What ideas can you add to _____? What would you suggest is a solution for _____?

•*Evaluation*

Do you agree with the statement that _____? Prioritize _____ according to _____. What criteria would you use to assess _____?

The questions at each level of cognitive complexity can vary in the demand they place on English language proficiency. For example, at the comprehension level students can name the main idea of the story either orally or in writing by selecting from among a number of alternatives or by generating the idea on their own. At the evaluation level, students can prioritize a list or provide criteria for the evaluation of an idea and write sentences indicating their assessment of a concept. Thus, *students can be assessed for their ability to respond to higher order questions even though they may have minimal skills in English.*

Scoring

Scoring criteria should focus on the level of student performance relative to a standard. The teacher can indicate student performance as being *effective, sometimes effective,* or *needs work,* as in Figure 2. Another option is to indicate a student's ability to perform at three levels: *independently, with assistance,* or *does not perform task* (Vygotsky, 1978). Below is an excerpt taken from a rating scale on reading comprehension. In this rating scale, a teacher could rate student performance in literal and inferential comprehension at the three performance levels as follows:

Reading Tasks	**Criteria**		
	Independently	*With assistance*	*Does not do task*
Understands literal meaning			
Draws inferences from reading			

Figure 2. Literacy Development Checklist

Student _____ **Teacher** _____

School _____ **Academic year** _____

MARK: [X] **Effective** [/] **Sometimes Effective** [–] **Needs work**

Reading Processes	*Quarter*			
	1	2	3	4
I. READING SKILLS				
Comprehends oral stories				
Reading vocabulary				
Fluent decoding				
Literal comprehension in reading				
Inferential comprehension				
II. INTEREST				
Initiates own reading				
Shows pleasure in reading				
Selects books independently				
Samples a variety of materials				
III. APPLICATIONS				
Participates in language experience story development				
Participates in reading discussion groups				
Writes appropriate dialogue journal entries				
Chooses books of appropriate difficulty				
Uses reading in written communication				
IV. READING STRATEGIES				
Monitors attention				
Notices miscues that interfere with meaning				
Infers meaning based on:				
Word clues				
Sentence structure				
Story structure				
Prior experience				
Summarizes main ideas or key events				
Links details to main ideas				
Remembers sequence of events				
Predicts conclusions				
Requests help if needed				

**Adapted from materials developed by the National Council of Teachers of English and by The Writing Lab of the University of New Hampshire.*

Writing Assessment

Purpose/Types

An assessment of writing provides an indication of how well students communicate in the written language. Teachers can determine student progress in writing through direct measures of functional literacy, such as writing samples, process writing, and dialogue journals. Writing samples are the most commonly used performance assessment of writing ability and will be described further.

Design

To construct a writing sample, the following steps are recommended:

(1) Select writing prompts that are developmentally appropriate. Use prompts that will elicit writing from students on topics that are familiar and interesting to them and that reflect writing styles commonly practiced at their grade level. Writing prompts should *not* depend upon a student's knowledge of a content area.

(2) Provide a choice of topics. By giving options, teachers increase the possibility that students will be interested in and capable enough to select one of the topics and write on it (Hamp-Lyons, 1990; Reid, 1990). However, because the scoring of writing samples may differ depending on the genre, the topics provided in a single assessment should be limited to one genre, preferably one to which students have already been exposed. Representative topics may differ by grade level and can include autobiographies, biographies, short stories, reports, and critical analyses.

(3) Check the topics for cultural bias. If the topics consistently assess knowledge of a particular culture rather than of the English language, they are probably culturally biased. Writing prompts may address relatively neutral themes, such as asking students to describe a favorite relative. Alternatively, they may include more controversial topics designed to provoke students to take a position, such as students' attitudes toward violence in school. Topics can be checked for bias with cultural informants (other school staff or parents) who share or have experience with the culture of the students.

Administration

Give clear directions. Students should know the amount of writing required, how much time they have to write, and how their writing will be scored. A minimum of thirty minutes should be provided to allow students to plan, organize, and revise their work. A minimum of two paragraphs should be required of students. Decide whether students can use dictionaries or other resource materials during the writing process and apply this rule consistently.

Scoring

In preparation for a writing assessment, students should be told what the scoring criteria are, be given model papers that illustrate the range of scores assigned to writing samples, and be provided opportunities through cooperative learning groups to discuss their work-in-progress relative to the scoring criteria (Kolls, 1992). Students should be given opportunities to re-write their products after receiving feedback from scoring.

Scoring writing samples of students in various grade levels means developing criteria to apply to different levels and determining what scores mean with regard to student progress. At a minimum, at least three categories should be established within each grade and level of English language proficiency with regard to writing ability: low, intermediate, and high levels of

writing. For example, a student might be a high level writer at an intermediate level of proficiency for sixth-graders.

To score a writing sample, use holistic scoring criteria which focus on the communicative nature of the writing (see **Figure 3**). Other options include assigning relative weights to overall fluency, organization, mechanics, and grammar. Depending on the goals of instruction, teachers may want to assign lower weights to language conventions (e.g., mechanics and spelling) and higher weights to expression (coherence, effective narrative strategies) in order to encourage students to communicate freely through writing.

The most reliable way to score writing samples is to allow at least one other teacher or staff member to score the sample separately, after both raters have discussed and agreed upon the scoring criteria and have obtained similar scores on a few writing samples. Raters need to establish consistently high interrater agreement levels with regard to the appropriate classification of each student but not necessarily with regard to exact scores. A minimum interrater agreement level of 80% is recommended. This means that two teachers rating a sample of student papers should agree on the overall rating for at least 80% of the students (see Figure 3).

Significant differences in the classification of students require negotiation, clarification of the scoring criteria, and, perhaps, re-scoring the writing sample. If differences persist, then raters should look at other evidence of student writing in order to determine progress and decide whether a student should be placed at a higher or lower level of instruction. The usefulness of the writing sample depends upon the quality of the scoring procedure used and the degree of interrater agreement established.

Anecdotal Records

Purpose

Anecdotal records are notes based on teacher observations of how students learn. They can be used to determine a student's integration of reading and writing processes. The open-ended nature of the anecdotal record allows teachers to describe in rich detail each student's development in literacy, to integrate these observations with other available information, and to identify instructional approaches that may be appropriate (Rhodes & Nathenson-Mejia, 1992).

Design

Anecdotal records can be produced by following three guidelines:

(1) describe a specific event, process, or product;

(2) report rather than evaluate—save interpretation for later; and

(3) relate the material to other known facts about the student.

Administration

Students can be observed while engaged in learning activities which produce the behaviors to be assessed. In completing the anecdotal record, the teacher will describe the specific learning event and what the student actually did in that situation. For example, teachers may develop anecdotal records of academic language skills by observing students as they work on math word problems in cooperative learning activities. Students may also be observed and asked appropriate questions while reading aloud in order to determine their level of reading comprehension on different types of materials. In addition to observations, teachers can review student products as a means of developing anecdotal records. For example, teachers can keep anecdotal records on student writing samples to document progress in literacy.

Figure 3. Sample Holistic Criteria

Rating	Criteria
5	• Vocabulary is precise, varied, and vivid • Organization is appropriate to writing assignment and contains clear introduction, development of ideas, and conclusion • Transition from one idea to another is smooth and provides reader with clear understanding that topic is changing • Meaning is conveyed effectively • A few mechanical errors may be present but do not disrupt communication • Shows a clear understanding of writing and topic development
4	• Vocabulary is adequate for grade level • Events are organized logically, but some part of the sample may not be fully developed • Some transition of ideas is evident • Meaning is conveyed but breaks down at times • Mechanical errors are present but do not disrupt communication • Shows a good understanding of writing and topic development
3	• Vocabulary is simple • Organization may be extremely simple or there may be evidence of disorganization • There are few transitional markers or repetitive transitional markers • Meaning is frequently not clear • Mechanical errors affect communication • Shows some understanding of writing and topic development
2	• Vocabulary is limited and repetitious • Sample is comprised of only a few disjointed sentences • No transitional markers • Meaning is unclear • Mechanical errors cause serious disruption in communication • Shows little evidence of discourse understanding
1	• Responds with a few isolated words • No complete sentences are written • No evidence of concepts of writing
0	• No response

Source: S. S. Moya, Evaluation Assistance Center (EAC)–East, Georgetown University, Washington, D.C., 1990.

Scoring/Interpretation

The interpretation of anecdotal records relies on teacher judgment rather than numerical scores. The teacher reflects on the meaning of the student's performance relative to instructional goals and activities and other information available on the student. Through anecdotal records, observations of students can be compared with information obtained from other assessment sources, such as reading comprehension tests and writing samples, and recommendations can be made for adapting instruction to student needs.

Self-Assessment: Language Proficiency

Purpose/Types

Self-assessment is an essential component of alternative assessment. Self-assessment enables students to reflect on their learning activities, task accomplishments, and progress in language development and other areas of instruction. Students see tangible evidence of learning when self-assessments are conducted periodically throughout the school year and can make plans to address areas where they feel they need more work. Students can also discuss their plans with their teacher and develop a schedule of learning and instructional activities through mutual agreement. Self-assessment measures of language proficiency can take the form of questionnaires, rating scales, and checklists.

Design

The limited research on self-assessment of language proficiency indicates that the more closely self-assessment measures relate to a student's reasons for using the language, the more reliable the information becomes (Oscarson, 1989). It has also been found that the type of questions used in self-assessments influence the size of the correlations between self-rating scores and scores obtained on language proficiency tests (Bachman & Palmer, 1989; LeBlanc & Painchaud, 1985).

Items that are related to the student's purposes for learning a language should provide more reliable information than those which are not. However, because the research on self-assessment of language proficiency has been conducted primarily with university students, the findings may have limited application at the upper elementary and middle grade levels. Teachers should, therefore, review the results of self-assessment relative to information generated from other sources.

One example of a rating scale for a self-assessment of reading ability is provided in **Figure 4**. Students are given a list of reading tasks and asked to indicate to what degree they think they can perform each task. Selected reading tasks should be appropriate to the students' grade level and reflect the local curriculum.

Figure 4. Self-Assessment of Reading Ability

In reading a passage, I can:

Reading Tasks	Criteria		
	All the time	*Sometimes*	*Almost never*
1. Understand the main idea			
2. Understand the details			
3. Understand the vocabulary			
4. Read quickly and still understand most of it			

Administration

One of the advantages of using self-assessment measures is that they can be administered individually, to groups of students, or to all students in the class simultaneously. They can be conducted at the student's leisure, perhaps on a bi-weekly basis, or whenever the teacher wants to get feedback from students on their learning progress.

Scoring

Scoring scales for self-assessment measures depend upon the types of questions used. A three-point scale is illustrated in Figure 4, although some scales consist of as many as five or ten points. Intervals on a five-point scale might be described as follows: (5 points) I can do this all the time; (4) I can do this most of the time; (3) I can do this about half the time; (2) I can seldom do this; and (1) I can never do this (LeBlanc & Painchaud, 1985).

Self-Assessment: Learning Strategies

Purpose

Many teachers either implicitly or explicitly ask students to use various learning strategies in performing classroom assignments. There are a number of planning, attentional, self-evaluative, and study skills that teachers may include in their curriculum and guide students in using (Ellis & Sinclair, 1989; O'Malley & Chamot, 1990; Oxford, 1990; Rubin & Thompson, 1982; Wenden & Rubin, 1987). If teachers want to know whether or not students are using these strategies, they can ask them to reflect on their use by performing a self-evaluation. Students can describe their strategies in dialogue journals or learning logs at the end of each unit. A *learning log* is an ongoing record kept by students to evaluate their own strategies for learning, although it can also be used to record language development and content knowledge. Learning logs enable students to identify goals they want to accomplish through learning activities.

Design

Students may be asked to indicate whether or not they used any special techniques to help themselves learn. The following strategies appear in Chamot, O'Malley, & Küpper (1991):

- I paid attention to the teacher.
- I took notes when I listened.
- I looked at the questions before I read.
- I took notes when I read.
- I looked at my notes after class.
- I repeated new words aloud.
- I used new words in sentences.

Administration and Scoring

Teachers can provide students with a weekly checklist to use for self-assessment and ask students to keep this information in a notebook. The checklist can also be placed in a student portfolio, and the teacher can add anecdotal notes regarding student progress in the systematic use of learning strategies.

Summary

Performance assessments are often used because they provide an opportunity to measure student performance across a range of skills and knowledge over time. Many performance measures are developed directly out of local curriculum objectives and student performance standards and, therefore, may be developed by teachers or other local staff. The performance measures suggested above can be used individually, in combination, or adapted to meet developmental and language proficiency needs of students. However, the use of multiple measures can lead to problems in recordkeeping and interpretation if teachers and other school staff lack ways of combining the information generated by them. In the following section, a framework for addressing these issues is presented.

Portfolio Assessment

Portfolios present a practical approach to assembling student work, interpreting evidence of student performance, and assessing student performance relative to instructional objectives. The concept of portfolios has been adopted from the arts where students maintain evidence of their best work to illustrate their accomplishments (Jongsma, 1989). In classroom instruction, portfolios are used in a similar manner, but the contents of the portfolio may represent work in progress, formal products, and ratings or other evidence of student knowledge relative to specific objectives or purposes (Valencia, 1990).

There is no "right" way to design portfolios. Each classroom, school district, and state will reflect a unique approach to authentic assessment, and in this sense, each student's collection of documents will differ somewhat, depending on the purpose of the assessment (Tierney, Carter, & Desai, 1991). Creating and maintaining student portfolios require that a variety of teacher and staff decisions be made concerning the instructional goals or objectives to be assessed, who will be involved in the portfolio design and interpretation, what instruments will be selected and how student performance will be demonstrated, how portfolio information will be used, and how the results will be conveyed to others. Because the entire portfolio process can be complex, systematic review and evaluation of the process should be conducted on a periodic basis.

Steps to portfolio development should include designing the portfolio, planning for and collecting the necessary data, analyzing the portfolio contents, and using the results (Moya & O'Malley, in press). Each of these points will be described in the following sections.

Designing Portfolios

For the purposes of assessment, the material in a student portfolio is most useful when each piece collected reflects progress toward particular learning goals. To this end, portfolios can be designed following a multi-step process that involves:

- setting the purpose of the portfolio;
- focusing on specific learning goals;
- identifying performance tasks and/or selecting appropriate instruments;
- setting criteria;
- selecting students to be assessed;
- collaborating with other teachers and staff;
- conducting staff development; and
- involving students and parents in the portfolio development process.

Each of these steps is discussed below.

Purpose

Before collecting any samples of student work, the first step in planning a portfolio is to determine the purpose for conducting the assessment, and how the results will be used (Moya & O'Malley, in press; Navarrete et al., 1990). Will the results be used for making decisions related to classroom instruction? Will they be used to determine whether a student is ready to move out of a special English language support program, such as ESL, ESL–content, or bilingual education? Will they be used to aid in assigning a student grade? Specifying how the results of the portfolio assessment are to be used will assist in determining the goals to be assessed and the samples of student work to be collected.

Specific Focus

The second step in portfolio design is focusing the portfolio on specific learning goals. Each portfolio should have a specific focus determined by school staff. The focus may be on oral or written language skills or on content area skills such as those required in mathematics, science, or social studies. Objectives may also be selected from goals contained in local curriculum frameworks, state guidelines, program objectives, or consensus among ESL/bilingual and mainstream teachers concerning important goals for learning. While it may be possible to collect student work for all content areas as well as for English language skills in a single student folder, attempting to do this for purposes of assessment could prove to be rather unwieldy. Providing a focus, on the other hand, avoids having to go through an overwhelming amount of information in each portfolio.

Performance Task/Instrument Selection

Once learning goals and performance objectives have been identified, portfolio designers need to identify performance tasks and instruments to be used to measure whether learning goals are being attained. School staff should strive to combine traditional and performance assessment measures in order to get multiple indicators of a student's ability level. Standardized tests are often required for district accountability needs. Using results obtained on standardized achievement tests together with anecdotal records, rating scales, teacher observation checklists, and writing samples to assess literacy skills provides much more information than standardized test results alone. Furthermore, having multiple indicators of student performance enables teachers to cross-check one type of information against another.

Each portfolio should also contain items which are required to assess progress on particular instructional goals and others which are optional. *Required* items might include those which are necessary to communicate a student's progress to other teachers or to administrators and can include a student's "best work," while *optional* items could include drafts of work in progress, ongoing ratings of performance, and occasional pieces selected by the student (Valencia, 1990). The use of required items introduces an element of consistency in the evaluation of student portfolios. By making certain items obligatory and others optional, teachers get the information they need for making instructional decisions while also encouraging students to participate actively in portfolio design and use.

Setting Criteria

Teachers or school staff should determine criteria (performance standards) for interpreting portfolio contents before collecting any student data. Performance criteria must be established in order to determine the degree to which a student has attained the objectives each task/instrument is designed to assess. Teachers need to identify and establish a minimum number of specific objectives that illustrate attainment of the instructional goals. One way to set criteria is to require students to perform tasks either independently or with assistance. Another possibility is to define expected student performance in narrative or anecdotal form. The narrative can specify what the students should be able to do to meet the criterion for performance or growth over time.

Staff Collaboration

If portfolio assessment is to be undertaken by a school-based team, it will be essential to identify school staff willing to participate in the assessment process. Ideally, a cross-section of teachers, staff, and administrators at each school who serve the same student(s) could become members of a portfolio assessment team. For example, a team at the upper elementary levels might consist of an ESL or bilingual education teacher, the grade level classroom teacher, a reading specialist, and the school principal. At the middle school level, the team might consist of a student's ESL or bilingual education teacher, content area teachers, and perhaps the school counselor. If portfolio assessment is a totally new experience for school staff, it is probably a good idea to pilot test the approach with a small number of staff and students before using it on a school-wide or district-wide basis.

Staff Development

All staff involved in the portfolio process should receive information and training on how to plan, implement, and interpret portfolios, especially when portfolio assessment is to be conducted at the school-building or district-wide level. Staff preparation not only enables staff to collaborate with and support each other, it also builds critical support for the portfolio process itself. Staff should receive training on how to design portfolios, how to target specific learning objectives and select students (if portfolios are limited to only a part of the student population), and how to set criteria for each portfolio. Staff development will also be essential to planning individual portfolio contents and to designing, administering, and scoring holistic, performance-based measures such as oral interviews, teacher observation checklists, rating scales, and writing samples. Performance and portfolio assessment hold great promise for improving assessment, but they can only reach their potential when teachers master their use (Stiggins, 1990).

Student Selection

Portfolio teams or individual teachers need to consider several factors when deciding whether to implement portfolio assessment with one or more students. If the classroom teacher is acting on his/her own to gather the information (without any support from other school staff), initially it may be advisable to limit the number of portfolios to only a few students. This can prevent teachers from being overwhelmed by the data collection and analysis effort and giving up before experiencing the benefits of portfolio assessment. On the other hand, if portfolio assessment is to be a school- or district-wide initiative, and if more than a few teachers are going to be involved and provided staff development in its systematic implementation, then many or all students can be included in the procedure.

Questions to ask regarding which students will participate in the portfolio process are: Will the portfolio be used in only one classroom? Will it be used only with students participating in a particular program (such as ESL, bilingual education, or Chapter 1)? Will the process be limited to a single or multiple grade levels? Will all students in each classroom be assessed using a portfolio? Will the procedure be limited to only those individuals needing frequent monitoring?

Student/Parent Involvement

The teacher and/or portfolio assessment team should encourage the active involvement of both students and parents in the assessment process. A key element in portfolio assessment is student self-evaluation. Students are asked to reflect on their progress toward learning goals and encouraged to select samples of their work which they believe illustrate progress toward these goals (Baron, 1992a; Palmer Wolf et al., 1992; Tierney, Carter, & Desai, 1991). Teacher/student/parent conferences can be scheduled at times convenient for the parents so that they can be informed of their child's progress. Portfolio contents provide much more information to parents about their child's learning than the percentiles represented on standardized achievement tests. Furthermore, parents appreciate being given an opportunity to respond to examples of student work, particularly items that show progress and substantiate grades given to the student.

Planning for and Collecting the Data

Data collection for portfolio assessment consists of identifying information to be entered into the portfolio, determining the frequency of data collection, creating a system to record and monitor the frequency of data collection, and setting guidelines for the removal and updating of portfolio material. These guidelines may vary depending on the purpose of the portfolio.

At the elementary and middle school levels, portfolios can be updated on a semester basis or at each grading period. Some schools are experimenting with portfolios which present increasingly comprehensive information on students by beginning with an initial portfolio every nine weeks, a semester portfolio drawn from these, and a year-end portfolio (Palmer Wolf et al., 1992). However, if the purpose is to closely monitor student progress, assessments should take place approximately every four to six weeks. Occasional items can be placed in the portfolio on a more frequent basis and removed when they have been superseded by more recent work or have become redundant. The decision to remove or maintain portfolio materials is a collaborative one to be made between the student and the teacher.

Analyzing Portfolio Contents

To determine whether a portfolio's contents reflect a student's progress toward learning goals, the teacher or portfolio assessment team can match contents to specific learning goals and objectives on a cover sheet, as illustrated in the Sample Portfolio Analysis Form in **Figure 5** (adapted by Pierce from Moya & O'Malley, in press). Note that student objectives are placed in the left-most column of Figure 5, followed by illustrations of student progress, and a specific citation or page reference to materials that support each objective. When additional evidence of student progress for each objective is entered or found in the portfolio, relevant page citations can be added.

Summary comments, interpretations, and recommendations can be added at the bottom of the Portfolio Analysis Form. These anecdotal notes help interpret and integrate the results of student performance across various measures or skill areas. Forms such as the one in Figure 5

Figure 5. Sample Portfolio Analysis Form

Student: Marisel A. Date: 5/1/92
Teacher: Jones Grade: 4
Educational Goal: Student demonstrates ability on a variety of writing tasks

Performance Task	Contents Illustrating Student Progress	Date
• Demonstrates interest and ability in a variety of writing	Literacy Development Checklist	3/20/92
• Writes a short story	Writing Sample: Dog Story	4/22/92
• Writes to communicate with others	Letter Dialog Journal	4/10/92 3/31/92
• Expresses writing preferences	Self-Assessment of Writing	4/24/92
• Shares writing with others	Anecdotal Record	4/6/92
Summary Comments:		

offer several advantages: teachers can indicate the relationship between each item in the portfolio and the objectives being assessed; they can make specific suggestions for instructional adaptations to student needs; and they can sift through portfolio contents periodically to remove materials that, although interesting, have no use in evaluating student progress.

To determine how students acquiring English as their second language are progressing in comparison to other students, performance assessments can be administered to native or proficient English-speaking grade-level peers in mainstream classes and average ratings calculated for these students as well as for the English language learners. Administering any of the performance assessments described in this publication to a locally selected, "average" group of English-speaking peers will provide the most meaningful basis for comparison. This will inform the teacher regarding both the English learner's progress and his/her preparation for functioning at an independent/average level in a typical mainstream classroom.

Using Portfolio Results

There are a variety of ways in which portfolio results can be used. The Sample Portfolio Analysis Form shown in Figure 5 is an essential component in many of these uses:

- *diagnosis and placement*—student strengths and needs are examined with regard to major curriculum objectives;
- *monitoring student progress*—growth in learning over the course of the semester or school year can be monitored;
- *feedback on the effectiveness of instruction*—if individual students are not progressing, the instructional approach should be re-evaluated and appropriate adaptations made to meet each

student's needs. One possible conclusion is that a student needs instructional support beyond the services provided by the classroom(s) in which the portfolio has been maintained;

- *communication with other teachers*—this includes other members of the portfolio team and those at other schools to which students may transfer;
- *student feedback*—portfolios enable students to comment and reflect on their progress and plan what they would like to do to maintain or change it; and
- *communication with parents*—portfolios provide parents with concrete evidence which supports instructional decisions.

Portfolios in Practice: Some Common Concerns

As a result of conducting workshops with teachers and administrators on portfolio assessment, it has been our experience that, while the use of portfolios has many advantages, it is not without its limitations. In this section we will describe some of the more commonly identified obstacles to implementing portfolios and suggest procedures for overcoming them. We will also look briefly at how portfolios have been used in some states and school districts and as part of ESL/bilingual programs to gain insights into real problems and possible solutions.

"Portfolios take too much time."

Teachers and administrators frequently express concerns that designing and using portfolios are time consuming processes. This is commonly expressed by teachers in the initial stages of portfolio assessment, but appears to become of lesser concern as teachers begin to establish a classroom environment that encourages students to be increasingly independent and responsible for assessing their own progress using portfolios (Kolls, 1992; Tierney, et al., 1991).

There are several ways to reduce the time involved in implementing portfolio assessment:

(1) make the data collection part of daily instructional routines;
(2) make students responsible for collecting information on a regular basis;
(3) identify specific items that go into the portfolio and list them on a portfolio analysis form;
(4) initially, use portfolios with only two or three students who need intensive monitoring;
(5) use staggered data collection cycles where assessment data are collected from only a few students daily or weekly;
(6) share responsibilities of data collection and interpretation with other school staff so that individual teachers do not become overwhelmed by the process; and
(7) create common planning times for teachers and other staff involved in portfolio development.

"How do we decide what goes into the portfolio?"

The teacher or portfolio assessment team determines the contents of each portfolio by identifying learning goals and specifying minimal levels of student performance that show whether students have attained these goals. Instruments are then selected for eliciting the desired student performance and criteria set for determining performance levels.

"How do we interpret the contents of a portfolio?"

A major concern in portfolio assessment is what to do with portfolio contents once they have been collected. Prior to attempting to interpret this information, the teacher or portfolio

assessment team should identify learning goals, student performance levels that demonstrate attainment of these goals, the tasks or instruments to be used to elicit student performance, the criteria for determining student progress, and whether specific portfolio items will be required or optional. Next, the teacher or assessment team must determine how much weight to give each item in the portfolio. As with the portfolio design process, guidelines for interpreting portfolio results should be established in advance of data collection.

"How are others using performance-based assessment and portfolios?"

Most of the portfolios in practice that we know of in language arts, ESL, and bilingual programs are used to assess literacy skills (see **Figure 6**). We know of few which focus on content area skills or oral language proficiency. In Figure 6 we provide a sampling of contents suggested for reading/writing portfolios by two ESL programs in two school districts in Virginia (Arlington and Fairfax) and two language arts programs in one county in Florida (Orange County) and one elementary school in New Hampshire (Stratham Memorial School).

What becomes immediately evident from Figure 6 is that while Arlington County and Stratham Memorial School distinguish between *reading* and *writing* entries, Orange and Fairfax Counties indicate *core* (required) and *optional* items, as we have recommended in this publication. In addition, all of the portfolios indicated in Figure 6 incorporate many of the items we have described in this publication, including test results, lists of books read, writing samples, learning logs, reading/writing checklists, and student self-assessments.

At the state level, more and more states are moving to performance assessment to ensure that students graduate with more than basic skills and to get a more complete picture of students' abilities. A 1990 survey by the Center for Research on Evaluation, Standards, and Student Testing (CRESST) at the University of California–Los Angeles found that nearly half of state testing programs either had performance assessments already in place, were planning to implement them, or were exploring their use (O'Neil, 1992). During 1990-91 the State of Vermont piloted the nation's first statewide portfolio assessment, which focused on mathematics and writing in Grades 4 and 8. The Vermont project was designed by teachers and fully implemented in 1991-92. It involves teachers in reviewing the actual classwork of students. It is important to note that major emphasis is being given to staff development and building resources that support portfolio assessment implementation efforts (Vermont State Department of Education, 1990).

In Connecticut, the State Department of Education has developed the Connecticut Common Core of Learning Assessment Project in Science and Mathematics to design performance-based assessment tasks for high school students that can be used by both teachers and policymakers to determine what students know and can do. This is a low-stakes project (the results are not being used for student promotion or graduation) which allows Connecticut educators time to examine their curricular, instructional, and assessment strategies and bring them into closer alignment with the new vision of science and mathematics education (Baron, 1992b).

In Kentucky, a state school reform mandate calls for the implementation of the nation's first completely performance-based statewide assessment system by 1995. The system, which has already begun to be implemented, will rely heavily on teacher assessment of student performance at Grades 4, 8, and 12. The statewide assessment system will have three components: an accountability assessment and two voluntary assessments—formal and informal—of student progress in classrooms. The state calls for teachers and other school staff to collect student performance task results in an "accountability portfolio" (Rothman, 1991).

Figure 6. Reading/Writing Portfolios: Sample Contents

A. Arlington County Public Schools, Virginia Elementary ESOL/HILT Program

Reading
- Teacher observation log
- Examples of what student can read
- Books/materials read
- Audiotape of student reading
- Test results, formal and informal
- Conferencing forms
- Examples of skills mastered

Writing
- First piece of writing each year
- Learning log, dialog journal
- January and May writing samples
- Drafts and final products from different genres (personal narratives, exposition, letters, poems, essays reports)
- Graphics (illustrations, diagrams)

B. Stratham Memorial (Elementary) School, New Hampshire Reading/Writing Portfolio

Reading
- Favorite books/authors list
- Genre graph, indicating type of literature preferred
- Journal entries
- List of completed books

Writing
- Writing sample and cover sheet
- List of completed pieces

Evaluation
- Goals and/or self-evaluation
- Annual narrative summary by student

C. Orange County Public Schools, Florida Literacy Portfolio Components

Core Elements
- Reading Development Checklist
- Three writing samples
- List of books read independently

Optional Elements
- Student self-assessment
- Reading journals
- Audiotapes of student reading
- "Things I Can Do" List
- Test results, formal and informal
- Reading comprehension tests
- Running records (miscue analysis) and anecdotal records

D. Fairfax County Public Schools, Virginia ESL Program

Core Elements
- Two writing samples
- Two oral production samples
- Informal reading assessment
- List of books read
- Results of Degrees of Reading Power (DRP) test, Grades 7–12

Optional Elements
- Dialogue journal excerpts
- Teacher observations
- Reading/writing checklists
- Student self-assessment
- Audio/videotapes
- Student-selected work

Summary and Conclusion

School-based student assessment procedures are beginning to change to keep pace with shifts in instructional paradigms. As instruction for language minority students has moved in the direction of teaching language and content-area skills in context, assessment has begun to incorporate a wider variety of measures that more closely reflect the types of tasks that students are asked to perform in classrooms or in real-life settings. In this sense, school-based assessment procedures are becoming more authentic. Changes in assessment are especially important for language minority students and those acquiring English as their second language.

In addition to shifts in school-based assessment practices, state and national education initiatives are moving toward higher standards of student performance and school accountability. Whether or not students and schools are prepared to meet standards will depend upon the nature of the standards, whether students are being assisted in meeting them, and what instruments are used to measure student progress. Clearly, these shifts toward higher performance standards will have an impact on language minority students.

To be able to effectively monitor the progress of language minority students, assessment needs to be conducted on an ongoing basis with procedures that promise to yield the most useful information for classroom instruction. While standardized achievement tests cannot provide this type of information, performance-based assessment can. Performance tasks and instruments must be carefully designed, administered, and scored by teachers or assessment teams who have been trained on the assessment procedures used. In addition, performance criteria need to be set and raters trained on appropriate scoring procedures to ensure the reliability and validity of results.

Portfolio assessment has the potential for becoming an effective management system for performance assessment results. It represents a focused assessment of learning goals or objectives in English, native language arts, and/or the content areas. It can be conducted on a classroom, school-wide, or district-wide basis. At all levels, specific steps are recommended for implementing portfolio assessment, including: identifying the purpose and focus of the assessment, designing the portfolio, collecting data, analyzing portfolio contents, and using analysis of the results to make adaptations in instruction.

In this publication, we have addressed how performance and portfolio assessment can be used to monitor the classroom progress of upper elementary and middle school students. Portfolio assessment has the potential for being used in other ways, as well, such as:

(1) For students who are moving from one teacher or school to another, portfolios can be used to pass along critical information on their strengths and needs so that the new teacher does not have to duplicate assessments which have already been conducted;

(2) For students who are being considered for placement at different levels within an ESL or bilingual education program, portfolio results can be used to determine their ability to function at various levels;

(3) For students who are being considered for transition from an ESL or bilingual education program to a mainstream, English-only program, portfolio results can be used to measure performance relative to grademates in the mainstream; and

(4) For students who are being considered for prereferral to special education programs, portfolio results can be used to determine whether performance is related to language proficiency, including both native language and English literacy skills.

Portfolio assessment is currently being used on a limited basis and at various levels of implementation with language minority students and English language learners in school districts and states throughout the country. A large number of teachers and staff, while valuing the nature of the information generated by student portfolios, nevertheless feel that the constraints placed on their time and resources by portfolio assessment preclude its use on a regular or expanded basis. We have discussed some of the concerns expressed by practitioners and have made a few suggestions for overcoming them. Key to addressing concerns about portfolios are an adequate program of staff development and opportunities for teacher collaboration in planning and collecting student performance samples. We hope that teachers and school staff continue to share with us their successes and failures in implementing portfolios as we move toward gathering more information on this innovative assessment approach.

References

Bachman, L.F., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing, 6*(1), 14–29.

Baron, J.B. (1992a, March). *Learning and assessment: Achieving the match.* Keynote speech presented at the National Forum on Assessment for the National Education Association, Washington, DC.

Baron, J.B. (1992b). *Performance-based assessment in mathematics and science: Examples from Connecticut's core of learning assessment program.* Paper presented at the National Forum on Assessment for the National Education Association, Washington, DC.

Brewer, W.R. (1990). *Performance assessment in Vermont.* (Policy Briefs Nos. 10 and 11.) Oak Brook, IL: North Central Regional Educational Laboratory.

Chamot, A.U., & O'Malley, J.M. (1987). The cognitive academic language learning approach (CALLA): A bridge to the mainstream. *TESOL Quarterly, 21*, 227–49.

Chamot, A.U., O'Malley, J.M., & Kupper, L. (1991). *Building bridges: Content and learning strategies for ESL* (Books 1–3). Boston, MA: Heinle & Heinle.

Cohen, A.D. (1980). *Testing language ability in the classroom.* Rowley, MA: Newbury House.

Council of Chief State School Officers (CCSSO). (1991). *State assessment of limited English proficient students.* Washington, DC: Author.

Cummins, J. (1982). *Tests, achievement, and bilingual students.* Wheaton, MD: National Clearinghouse for Bilingual Education.

Cummins, J. (1983). Conceptual and linguistic foundations of language assessment. In S.S. Seidner (Ed.), *Issues of language assessment: Language assessment and curriculum planning.* Wheaton, MD: National Clearinghouse for Bilingual Education.

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy.* Clevedon, U.K.: Multilingual Matters.

Cunningham, P.M., Moore, S.A., Cunningham, J.W., & Moore, D.W. (1983). *Reading in elementary classrooms: Strategies and observations.* White Plains, NY: Longman.

DeGeorge, G.P. (1988). *Assessment and placement of language minority students: Procedures for mainstreaming.* Silver Spring, MD: National Clearinghouse for Bilingual Education.

Development Associates. (1987). *Student oral proficiency rating.* Arlington, VA: Author.

Dole, J.A., Duffy, G.G., Roehler, L.R., & Pearson, P.D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research, 61*, 239–64.

Ellis, G., & Sinclair, B. (1989). *Learning to learn English.* Cambridge, U.K.: Cambridge University Press.

Freeman, Y.S., & Freeman, D.E. (1992a, Winter/Spring). *Portfolio assessment for bilingual learners.* Alexandria, VA: TESOL Bilingual Education Interest Section.

Freeman, Y.S., & Freeman, D.E. (1992b). *Whole language for second language learners*. Portsmouth, NH: Heinemann.

Genishi, C. (1985). Observing communicative performance in young children. In A. Jagger and M.T. Smith-Burke (Eds.), *Observing the language learner* (pp. 131–142). Newark, DE: International Reading Association.

Gonzalez Pino, B. (1988, September). *Testing second language speaking: Practical approaches to oral testing in large classes*. New York: Northeast Conference of Teachers of Foreign Languages.

Goodman, Y.M., & Burke, C.L. (1972). *Reading miscue inventory: Procedures for diagnosis and evaluation*. New York: Macmillan.

Goodman, Y.M., Watson, D.J., & Burke, C.L. (1987). *Reading miscue inventory: Alternative procedures*. New York: Richard C. Owen.

Goodman, K.S., Goodman, Y.M., & Hood, W.J. (Eds.). (1989). *The whole language evaluation book*. Portsmouth, NH: Heinemann.

Graves, D.H. (1983). *Writing: Teachers and children at work*. Portsmouth, NH: Heinemann.

Hamayan, E.V., Kwiat, J.A., & Perlman, R. (1985). *The identification and assessment of language minority students: A handbook for educators*. Arlington Heights, IL: Illinois Resource Center.

Hamp-Lyons, L. (1991). *Assessing second language writing in academic contexts*. Norwood, NJ: Ablex Publishing Co.

Hamp-Lyons, L. (1990). Second language writing: Assessment issues. In B. Kroll (Ed.), *Second language writing*. New York: Cambridge University Press.

Haney, W., & Madaus, G. (1989). Searching for alternatives to standardized tests: Whys, whats, and whithers. *Phi Delta Kappan, 70*, 683–87.

Herman, J.L. (1992, May). What research tells us about good assessment. *Educational Leadership, 49*(8), 74–78.

Jongsma, K.S. (1989). Portfolio assessment. *The Reading Teacher, 43*, 264–65.

Kolls, M.R. (1992, March). *Portfolio assessment: A feasibility study*. Paper presented at the annual meeting of the Teachers of English to Speakers of Other Languages. Vancouver, British Columbia, Canada.

LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly, 19(4)*, 673–687.

Maryland State Department of Education. (n.d.). *Questions for quality thinking*. Baltimore, MD: Author.

Mitchell, R. (1992). *Testing for learning: How new approaches to evaluation can improve American schools*. New York: The Free Press.

Mohan, B.A. (1986). *Language and content*. Reading, MA: Addison-Wesley.

Moya, S.S., & O'Malley, J.M. (in press). A portfolio assessment model for ESL. *Journal of Educational Issues of Language Minority Students.*

Navarrete, C., Wilde, J., Nelson, C., Martinez, R., & Hargett, G. (1990). *Informal assessment in educational evaluation: Implications for bilingual education programs.* Washington, DC: National Clearinghouse for Bilingual Education.

Neill, D.M., & Medina, N.J. (1989). Standardized testing: Harmful to educational health. *Phi Delta Kappan, 70,* 688–97.

Oller, J.W. (1979). *Language tests at school.* London: Longman.

Omaggio, A.C. (1986). *Teaching language in context.* Boston, MA: Heinle & Heinle.

O'Malley, J.M. (1989). Language proficiency testing with limited English proficient students. In J.E. Alatis (Ed.), *Georgetown University Roundtable on Languages and Linguistics* (pp. 235–44). Washington, DC: Georgetown University.

O'Malley, J.M. (1991, September). *Looking for academic language proficiency: Comments on Damico's paper.* Paper presented at the Second National Research Symposium on LEP Student Issues, Washington, DC.

O'Malley, J.M., & Chamot, A.U. (1990). *Learning strategies in second language acquisition.* Cambridge, U.K.: Cambridge University Press.

O'Malley, J.M., & Pierce, L.V. (1991). *A survey of state assessment practices affecting language minority students.* Arlington, VA: Georgetown University Evaluation Assistance Center (EAC)–East.

O'Malley, J.M., & Pierce, L.V. (1992, March). *Portfolio assessment: Experiences from the field.* Paper presented at the annual meeting of the Teachers of English to Speakers of Other Languages. Vancouver, British Columbia, Canada.

O'Neil, J. (1992, May). Putting performance assessment to the test. *Educational Leadership, 49*(8), 14–19.

Oscarson, M. (1989). Self-assessment of language proficiency; rationale and applications. *Language Testing, 6*(1), 1–13.

Oxford, R.L. (1990). *Language learning strategies: What every teacher should know.* New York: Newbury House.

Padron, Y.N., & Waxman, H.C. (1988). The effects of ESL students' perceptions of their cognitive strategies on reading achievement. *TESOL Quarterly, 22,* 146–50.

Palmer Wolf, D., LeMahieu, P.G., & Eresh, J. (1992, May). Good measure: Assessment as a tool for educational reform. *Educational Leadership, 49*(8), 8–13.

Pinnell, G.S. (1985). Ways to look at the functions of children's language. In A. Jaggar & M.T. Smith-Burke (Eds.), *Observing the language learner* (pp. 57–72). Urbana, IL: National Council of Teachers of English and International Reading Association.

Pressley, M., et al. (1990). *Cognitive strategy instruction that really improves children's performance.* Cambridge, MA: Brookline Books.

Reid, J. (1990). Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective. In B. Kroll (Ed.), *Second language writing*. New York: Cambridge University Press.

Rhodes, L.K., & Nathenson-Mahia, S. (1992). Anecdotal records: A powerful tool for ongoing literacy assessment. *The Reading Teacher, 45*, 502–09.

Roeber, E.D. (1990). *Performance assessment: A national perspective* (Policy Briefs Nos. 10 and 11). Oak Brook, IL: North Central Regional Educational Laboratory.

Rothman, R. (1991, May 8). Ambitious student assessment system advances. *Education Week*, p. 14.

Rubin, J., & Thompson, I. (1982). *How to be a more successful language learner.* Boston, MA: Heinle & Heinle.

Saville-Troike, M. (1991). *Teaching and testing for academic achievement: The role of language development.* Washington, DC: National Clearinghouse for Bilingual Education.

Spanos, G., Rhodes, N.C., Dale, T.C., & Crandall, J. (1988). Linguistic features of mathematical problem solving: Insights and applications. In R.R. Cocking & J.P. Mestre (Eds.), *Linguistic and cultural influences on mathematics* (pp. 221–240). Hillsdale, NJ: Lawrence Erlbaum.

Stiggins, R.J. (1990). *The foundation of performance assessment: A strong training program* (Policy Briefs Nos. 10 and 11). Oak Brook, IL: North Central Regional Educational Laboratory.

Stiggins, R.J. (1987). Design and development of performance assessments. *Educational Measurement: Issues and Practices, 6*(1), 33–42.

Tierney, R.J., Carter, M.A., & Desai, L.E. (1991). *Portfolio assessment in the reading–writing classroom.* Norwood, MA: Christopher Gordon Publishers.

Ulibarri, D.M., Spencer, M.L., & Guillermo, A.R. (1981). Language proficiency and academic achievement: A study of language proficiency tests and their relationship to school ratings as predictors of academic achievement. *Journal of the National Association for Bilingual Education, 5*(3), 47–80.

Valencia, S.W. (1990). Alternative assessment: Separating the wheat from the chaff. *The Reading Teacher, 43*, 60–61.

Vermont State Department of Education. (1990). *Vermont writing assessment: The pilot year.* Montpelier, VT: Author.

Vygotsky, L.M. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Wenden, A., & Rubin, J. (1987). *Learner strategies in language learning.* Englewood Cliffs, NJ: Prentice-Hall.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan, 70*, 703–13.

Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership, 46*(7), 41–47.

About the Authors

Lorraine Valdez Pierce is a Research Associate for the Center for Applied Linguistics and Coordinator of Publications for NCBE. She has been Senior Evaluation Specialist at the Georgetown University Evaluation Assistance Center (EAC)–East, Director of National Origin Desegregation Programs at American University's Mid-Atlantic Equity Center, and an ESL teacher of adolescents and adults for over a decade. Her latest publication is *Effective Schools for National Origin Language Minority Students*.

J. Michael O'Malley is Supervisor of Testing and Accreditation at Prince William County Public Schools in Manassas, Virginia. He was Director of the Georgetown University Evaluation Assistance Center (EAC)–East from 1985 to 1991 and has served as the Principal Investigator for Georgetown University's Language Research Projects on language learning strategies. His latest publication (with A. U. Chamot) is *Learning Strategies in Second Language Acquisition*.

ED/OBEMLA92-3

BEST COPY AVAILABLE